Title
Physical symmetries and Fisher's information measure

Journal
Physics Letters A, 373

Author
Soffer, Bernard H.

Publication Date
2009

DOI
10.1016/j.physleta.2009.01.007

Peer reviewed

Physical symmetries and Fisher's information measure

F. Pennini, [a,b] A. Plastino, [a] B.H. Soffer, [c,*] C.Vignat [d]

[a] *Facultad de Ciencias Exactas, Universidad Nacional de La Plata Argentina's National Research Council (IFLP-CCT-CONICET), C.C. 727, (1900) La Plata, Argentina*

[b] *Departamento de Física, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile.*

[c] *University of California at Los Angeles, CA 90095, USA*

[d] *I.G.M., Université de Marne la Vallée*

Abstract

We investigate the physical consequences of imposing symmetry requirements to the Cramer-Rao inequality, and investigate in particular translation, inversion, and rotation.

Key words: Fisher Information, invariance, symmetries, Cramer-Rao
PACS: 2.50.-r, 3.67.-a, 05.30.-d

1 Introduction

Fisher's information measure is much in vogue nowadays in the physics world and many authors are engaged in Projects that revolve around it [1]. The most complete exposition of Fisher's measure and its scientific applications can be found in the the books by Frieden [1]. However, not much attention is paid therein to the ordinary Lie symmetries that are very common throughout Physics. We start addressing the issue in this effort, by studying the effects of the most common symmetries on the Cramer-Rao (CR) inequality, a cornerstone of the subject, fundamentally based in measurement theory.

* Corresponding author.
Email addresses: fpennini@ucn.cl, pennini@fisica.unlp.edu.ar (F. Pennini,), plastino@fisica.unlp.edu.ar (A. Plastino,), bsoffer@ucla.edu (B.H. Soffer,), vignat@univ-mlv.fr (C.Vignat).

Why should this be of interest? Because the CR is regarded as reflecting experimental laboratory scenarios. The ensuing empirical results referring to the statistical dispersion (variance) of results can not depend upon say, a rotation of our instruments. Thus, the CR must be invariant under elementary transformations. This simple issue acquire momentous importance when we realize that most physical theories can be derived from a spacial variational principle, called the *extreme physical* (i.e., Fisher's) *information* one (EPI) [1] that is precisely based upon the CR relations [1].

Our results should be of interest to the many Fisher practitioners of the physics-community. Let us insist in that our motivation is simple enough: using a given coordinate system involves always a choice. In a physical relationship like the Cramer-Rao inequality [1], actual numerical physical results should not depend on it.

2 Shift-invariant Fisher measure

2.1 One dimensional instance

Sir R.A. Fisher advanced in the 1920s an extraordinarily important scientific tool: the measure of information that bears his name (for details and discussions refer to [1–3]). Fisher information (FI) arises as a measure of the expected error in a measurement [1]. A particular FI-case of great practical importance is that of translation families [1,4], i.e., distribution functions whose *form* does not change under displacements of a shift parameter θ. The ensuing distributions are shift-invariant –à la Mach, with no absolute origin for θ– and where the ensuing FI obeys Galilean invariance [1]. The next Section clarifies the concept in a somewhat more elaborate fashion.

Using a more precise notation than ordinarily employed, our interest will always lie in the random variable X that takes values η in the range $[a,b]$ with probability $\int_a^b d\eta f_X(\eta)$. As a consequence, the classical Fisher information associated with translations of a one-dimensional observable x with probability density $f(x)$ becomes [5]

$$I_X = \int \mathrm{d}x\, f_X(x) \left[\frac{\partial \ln f_X(x)}{\partial x}\right]^2 , \tag{1}$$

and the associated Cramer–Rao inequality [1,5] satisfies the relation

$$I_X\, \Delta X^2 \geq 1 \tag{2}$$

where ΔX^2 is the variance of the stochastic variable X [5]

$$\Delta X^2 = \int \mathrm{d}x\, f_X(x)\, x^2 - \left(\int \mathrm{d}x\, f_X(x)\, x\right)^2, \tag{3}$$

and represents the mean-square error associated to the pertinent measurement.

We will address the issue of the Lie symmetries of the Cramer Rao relation (CR), beginning with those most common in Physics: The Galilean symmetries of scale and translation, Inversion, and Rotation. We prove below that the CR relation is invariant under the concomitant transformations, with a surprising non-trivial caveat. We give a particular illustration for the Harmonic Oscillator.

2.2 The multivariate case

Let us consider now the multivariate instances [5]. We deal of course with

$$\mathbf{X} \in \mathcal{R}^n. \tag{4}$$

It is known that the $n \times n-$Fisher matrix becomes [5]

$$\mathbb{I}_X = \int \mathrm{d}\mathbf{x}\, f_X(\mathbf{x})\, [\nabla \ln f_X(\mathbf{x})][\nabla \ln f_X(\mathbf{x})]^\top, \tag{5}$$

where ∇ is the $n-$gradient operator and $\top$ denotes transposition [5].The covariance matrix of $\mathbf{X}$ is

$$\mathrm{Cov}(\mathbf{X}) \equiv \mathbb{K}_X = \langle \mathbf{X}\mathbf{X}^\top \rangle - \langle \mathbf{X} \rangle \langle \mathbf{X}^\top \rangle. \tag{6}$$

Then, the Cramer–Rao inequality that generalizes to an inequality (2) that has the following aspect [5] ($\mathbb{I}_n$ is the identity matrix)

$$\mathbb{I}_X\, \mathrm{Cov}(\mathbf{X}) = \mathbb{I}_X\, \mathbb{K}_X \geq \mathbb{I}_n. \tag{7}$$

3 General result

If $\mathbb{A}$ is a scaling matrix of determinant $det(\mathbb{A}) \neq 0$ one has [6]

$$\mathbb{I}_{AX} = [(\mathbb{A})^{-1}]^\top \mathbb{I}_X (\mathbb{A})^{-1}, \tag{8}$$

$$\mathbb{K}_{AX} = \mathbb{A}\mathbb{K}_X(\mathbb{A})^\top, \tag{9}$$

where this scaling matrix could well be a rotation matrix of determinant $det(\tilde{A}) = 1$ [8]. Now,

$$\begin{aligned} \mathbb{I}_{AX}\mathbb{K}_{AX} &= [(\mathbb{A})^{-1}]^\top \mathbb{I}_X(\mathbb{A})^{-1}\mathbb{A}\mathbb{K}_X(\mathbb{A})^\top = \\ &= [(\mathbb{A})^{-1}]^\top \mathbb{I}_X\mathbb{K}_X(\mathbb{A})^\top \geq [(\mathbb{A})^{-1}]^\top(\mathbb{A})^\top \geq \mathbb{I}_n. \end{aligned} \tag{10}$$

entailing, from a physicist viewpoint, that the CR keeps its form after the transformation. Of course, a mathematician would observe that, since the two members of such transformation are themselves random vectors, the Cramer-Rao inequality holds for both of them. Most physicists would probably prefer the less abstract proof that we also provided. Notice that no invariance can be claimed according to the result (10). However, if the CR-bound is saturated, in the sense that the equality holds

$$\mathbb{I}_X\mathbb{K}_X = \mathbb{I}_n, \tag{11}$$

then, since it is an elementary matrix property that $[(\mathbb{A})^{-1}]^\top = [(\mathbb{A})^\top]^{-1}$, (10) immediately yields

$$\mathbb{I}_{AX}\mathbb{K}_{AX} = \mathbb{I}_X\mathbb{K}_X = \mathbb{I}_n. \tag{12}$$

If the lower bound is actually attained, the situation persists under the change $X \to \mathbb{A}X$.

Note that the CR bound is saturated in the Gaussian case (actually, only in this situation). Since in this instance $\mathbb{A}\mathbb{X}$ is Gaussian too (provided $\mathbb{A}$ is invertible), this makes the result straightforward. In the quantum realm this "Gaussianity" is associated to coherent states (Gaussian wave packets), which actually represent those vectors in Hilbert's space that most closely resemble classical states [9]. They "saturate" the uncertainty principle, closely related to the CR [1]. We conjecture that CR-saturation is somehow an indicator of classicity (except for the ground state of the harmonic oscillator, the only exception to the above remark).

4 Elementary symmetries

(1) *Scaling*

We consider here the most common symmetries. For a very interesting treatment of this subject in relation to other information measures, although made in a different, nonphysical context, the reader is referred to [7].

If a change of units is performed so that $X \rightarrow aX$, a a non-null real parameter. With $Z = aX$,

$$\int_a^b duF_Z(u) = \int_a^b dvF_X(v) = 1, \tag{13}$$

entailing

$$f_Z(u) = \frac{f_X(u/a)}{|a|}.$$

Therefore, the Fisher measure changes as follows

$$I_{aX} = \int \mathrm{d}u\, f_Z(u) \left[\frac{\partial \ln f_Z(u)}{\partial u}\right]^2, \tag{14}$$

and the variance is

$$\Delta(aX)^2 = a^2 \Delta X^2. \tag{15}$$

Since

$$\frac{\partial \ln f_Z}{\partial u} = \frac{1}{a}\frac{f_X'(u/a)}{f_X(u/a)}, \tag{16}$$

one immediately ascertains that

$$I_{aX} = \frac{I_X}{|a|^2} \tag{17}$$

then, the Cramer-Rao inequality has the form

$$I_{aX}\, \Delta(aX)^2 = I_X \Delta X^2 \geq 1, \tag{18}$$

using the inequality (2). Cramer-Rao is scale invariant, as it should be, since it cannot be affected by a change in units, say, from centimeters to meters. Note also that taking traces in Eq. (10) one finds

$$Tr[\mathbb{I}_{AX}\mathbb{K}_{AX}] = Tr[\mathbb{I}_X \mathbb{K}_X] \geq n, \tag{19}$$

i.e., a degenerate trace version of the CR inequality, obviously invariant for any linear transformation generated by $\mathbb{A}$. As far as we know this is a new result. In one dimension it trivially entails invariance of the CR under scaling.

(2) *Inversion*: We need to effect the change $X \rightarrow -X$ and $f_X(x) \rightarrow f_{-X}(x)$. Using (15-17) with $a = -1$ yields then

$$I_X = I_{-X};\ \textit{and}\ \Delta(-X)^2 = \Delta X^2, \tag{20}$$

so that the CR remains invariant under inversion.

(3) *Translation*: we change $X \rightarrow Z = X + \epsilon$, with ϵ a real parameter. In this case, the prevailing scenario, with normalization according to (13), leads to

$$f_Z(z) = f_X(z-\epsilon); \;\; z = x+\epsilon; \;\; dz = dx.$$

and the Fisher measure is

$$I_{Z=X+\epsilon} = \int \mathrm{d}z\, f_Z(z-\epsilon) \left[\frac{\partial \ln f_Z(z-\epsilon)}{\partial z}\right]^2, \tag{21}$$

which clearly implies $I_Z = I_X$, while the variance is

$$\Delta(X+\epsilon)^2 = \Delta X^2, \tag{22}$$

so that

$$I_{X+\epsilon} = I_X \tag{23}$$

and the Cramer-Rao relation is

$$I_{X+\epsilon}\, \Delta(X+\epsilon)^2 = I_X\, \Delta X^2 \geq 1. \tag{24}$$

The above invariance properties clearly reflect just what the expression shift-invariance, invoked in defining shape invariant families of probability distributions, actually means.

4.1 Odd situations

The linear and Hermitian parity operator Π has eigenvalues ± 1. Let $\psi_\pm$ be eigenfunctions of Π. Thus, within Hilbert's space $\mathcal{H}$, any $\psi(x) \in \mathcal{H}$ writes

$$\psi(x) = a_+\psi_+(x) + a_-\psi_-(x)\,;\; \psi^*(x) = a_+^*\psi_+^*(x) + a_-^*\psi_-^*(x) \;\; ;\; a_+^*a_+ + a_-^*a_- = 1, \tag{25}$$

and the probability density

$$P(x) = (\psi^*\psi)(x) = |a_+|^2|\psi_+|^2 + |a_-|^2|\psi_-|^2 + a_+^*a_-\psi_+^*\psi_- + C.C., \tag{26}$$

has both even and odd parity-components, so that it can be cast in the form

$$P(x) = C_+P_+(x) + C_-P_-(x). \tag{27}$$

We wonder what happens to the odd-part, which should be 'smaller' than the even part so as not to lead to a negative PDF under sign inversion of the argument. Since 1, x, x^2, x^3, ... is a basis, any function $f(x) \in \mathcal{H}$ has

$$\langle M_{ij} \rangle = \int \mathrm{d}\Gamma \, \rho(\mathbf{z}) \, M_{ij}. \tag{31}$$

We appreciate the fact that $I(f)$ is a $3N \times 3N-$dimensional matrix.

We shall here consider a simpler classical system governed by a time-independent Hamiltonian $\mathcal{H}$ with just two degrees of freedom (one "generalized coordinate" x and one "generalized momentum" p). In this case the "state-vector" will be $\mathbf{z} \equiv (x, p)$.

We presuppose that the system is in thermal equilibrium at temperature T, and that it is appropriately described by Gibbs' canonical probability distribution. All our results below depend upon this assumption. At point $\mathbf{z}$, this canonical ensemble probability density $\rho(\mathbf{z})$ reads

$$\rho(\mathbf{z}) = \frac{e^{-\beta \mathcal{H}(\mathbf{z})}}{Z} \tag{32}$$

with $\beta = 1/kT$ and k Boltzmann's constant. If h denotes an elementary cell in phase-space, we write, with some abuse of notation [10], $\mathrm{d}\Gamma = \mathrm{d}x \, \mathrm{d}p / 2\pi\hbar$ for the pertinent integration-measure. Notice that the only role of $\hbar$ here is that of balancing dimensions. Our calculations are entirely classical ones and we start with the partition function whose form is [10]

$$Z = \int \mathrm{d}\Gamma \, e^{-\beta \mathcal{H}(\mathbf{z})}. \tag{33}$$

It is easy to see that (for details and discussions, see please Ref. [11])

$$\nabla_p \ln \rho(\mathbf{z}) = -\beta \, \nabla_p \mathcal{H}(\mathbf{z}) \tag{34a}$$
$$\nabla_x \ln \rho(\mathbf{z}) = -\beta \, \nabla_x \mathcal{H}(\mathbf{z}). \tag{34b}$$

The $2 \times 2-$matrix (29) has then the form

$$M_{ij} = \beta^2 \begin{pmatrix} (\nabla_x \mathcal{H})^2 & \nabla_x \mathcal{H} \nabla_p \mathcal{H} \\ \nabla_x \mathcal{H} \nabla_p \mathcal{H} & (\nabla_p \mathcal{H})^2 \end{pmatrix} \tag{35}$$

and the Fisher information matrix adopts the appearance

$$I(f) = \beta^2 \begin{pmatrix} \langle (\nabla_x \mathcal{H})^2 \rangle & \langle \nabla_x \mathcal{H} \nabla_p \mathcal{H} \rangle \\ \langle \nabla_x \mathcal{H} \nabla_p \mathcal{H} \rangle & \langle (\nabla_p \mathcal{H})^2 \rangle \end{pmatrix}. \tag{36}$$

We can compute explicitly the determinant of the matrix $I(f)$ arriving at

$$|I(f)| = \beta^2 \left\{\langle(\nabla_x \mathcal{H})^2\rangle\langle(\nabla_p \mathcal{H})^2\rangle - \langle\nabla_x \mathcal{H}\nabla_p \mathcal{H}\rangle^2\right\}. \tag{37}$$

Eqs. (36)-(37) are quite general expressions for the classical Fisher information.

5.1 Harmonic oscillator

For a concrete example we turn to the very important (and usual test case) of the the harmonic oscillator. Its classical Hamiltonian is given by $\mathcal{H} = p^2/2m + m\omega^2 x^2/2$. Given that in such an instance the canonical ensemble probability distribution becomes Gaussian, the CR-relation is saturated and rotational invariance automatically ensues. Notice that we are here speaking of *rotational invariance in the* $x - p$ *phase-space plane*, of obvious relevance for the theory of canonical transformations.

6 Conclusions

After a rather general study of the invariance-properties associated to the Cramer-Rao bound, our results can be summarized as follows

- We have verified that the Cramer-Rao relation is invariant under translation, inversion, and scaling.
- This adds new meaning to the expression shift-invariant Fisher measure.
- If the CR-bound is saturated, it remains so under any transformation represented by a square matrix, representative of a (physical) Unitary operator.
- Such is the situation, for example, in the special case of the statistical treatment of the classical harmonic oscillator.
- Quite general expressions for the classical Fisher information in phase-space are given.

Acknowledgment F. Pennini would like to thank partial financial support by FONDECYT, grant 1080487.

References

[1] B.R. Frieden, Physics from Fisher information (Cambridge University Press, Cambridge, England, 1998); B.R.Frieden and Robert A. Gatenby (Eds.) Exploratory Data Analysis Using Fisher Information (Springer-Verlag, London,

2007); R. Frieden, Science from Fisher information (Cambridge University Press, Cambridge, England, 2004).

[2] B.R. Frieden, B.H. Soffer, Phys. Rev. E **52** (1995) 2274.

[3] F. Pennini and A. Plastino, Phys. Rev. E **69** (2004) 057101.

[4] F. Pennini, A.R. Plastino, and A. Plastino, Physica A **258** (1998) 446.

[5] Michael J.W Hall, Phys. Rev. A **62** (2000) 012107.

[6] R. Zamir, Tech. Rep. EE-S-97-11, Tel Aviv University (1997).

[7] L. Knockaert, Information Sciences **152** (2003) 139.

[8] J. J. Sakurai, Modern quantum mechanics (Benjamin, Amsterdam, 1985).

[9] A. Peres, Quantum theory: concepts and methods (Kluer, London, 1993).

[10] R.K. Pathria, *Statistical Mechanics* (Pergamon Press, Exeter, 1993).

[11] F. Pennini, A. Plastino, *Physics Letters A* **349** (2006) 15.

Note added in
proof.
Its worth
mentioning that

Note added in
proof.
Its worth
mentioning that

Note added in proof.
Its worth mentioning that our CI and Fisher information [FI] symmetry results could have been inferred by an interpretation of the general symmetry principle, that an effect is at least as symmetric as its cause. Scientific laws express particular causal relations. The data including its symmetry that the CI and concomitant FI laws receive would represent the physical cause, and the results that those laws give would represent the physical effect [16]. We havepreferred to give a detailed,far less abstract and more accessible development in this paper.

J. Rosen, Symmetry Rules: How Science and Nature are Founded on Symmetry, Springer-Verlag, Heidelberg, Berlin, 2008

Note added in proof.
Its worth mentioning that our CI and Fisher information [FI] symmetry results could have been inferred by an interpretation of the general symmetry principle, that an effect is at least as symmetric as its cause. Scientific laws express particular causal relations. The data including its symmetry that the CI and concomitant FI laws receive would represent the physical cause, and the results that those laws give would represent the physical effect [16]. We havepreferred to give a detailed,far less abstract and more accessible development in this paper.

[16] J. Rosen, Symmetry Rules: How Science and Nature are Founded on Symmetry, Springer-Verlag, Heidelberg, Berlin, 2008